Prospects for b-tagging with ATLAS and tracking results with cosmics

20th Hadron Collider Physics Symposium
Évian

Laurent VACAVANT
(on behalf of the ATLAS Collaboration)

Motivation and prospects

B-tagging key tool for: HF production measurements (bb, tt, …), NP searches, low-mass Higgs boson searches, etc

One more incentive: S/B less favorable @ 7 TeV for top rediscovery in tt pairs:

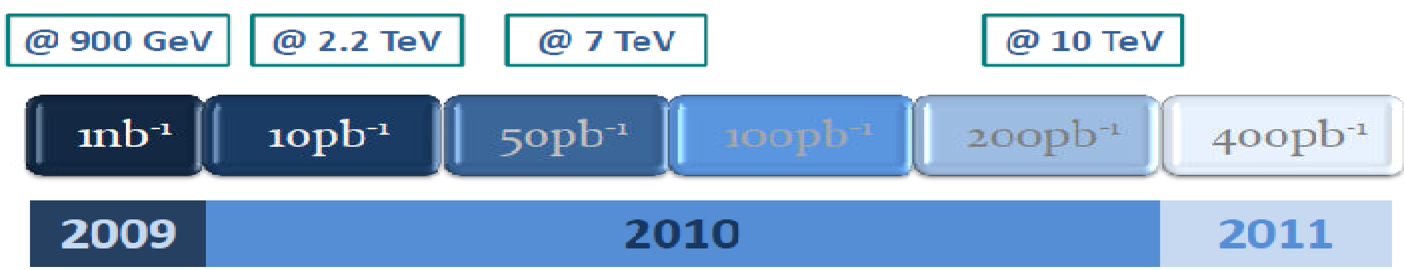
Selected events (no b-tagging):

eνb jjb	$200pb^{-1}$ @10TeV	$20pb^{-1}$ @7TeV
ttbar	1286	51
W+jets	448	28
Single top	81	4
Other bkgd	67	7
S/B	**2.1**	**1.3**
$S/\sqrt{(S+B)}$	30	5.4

Back-of-the-enveloppe estimate requiring one b-tagged jet:

$20pb^{-1}$ @7TeV	No b-tag	1 b-tag : JP60
ttbar	51	~35
W+jets	28	<1
S/B	**1.3**	**2.9**

Giving us motivation for an aggressive roadmap:



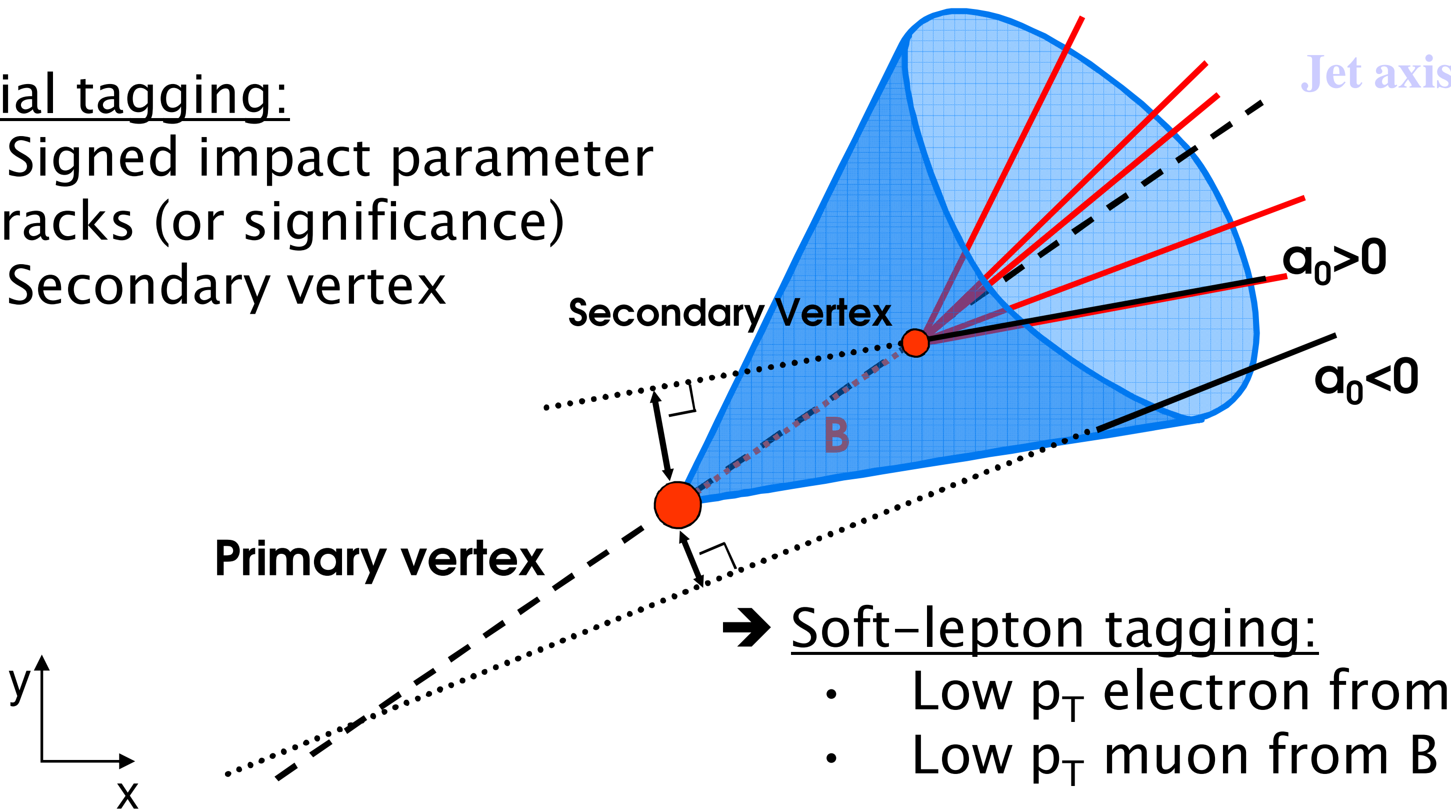
- tracking/vertexing commissioning, MC validation
- early taggers (on light jets : 30k (<300 b-jets))
- calibration with dijets
- bb production
- new physics
- top rediscovery
- top cross-section w b-tagging
- sophisticated taggers
- calibration with ttbar

b-tagging basics

Hard fragmentation of b quarks $x_B \sim 70\%$
High mass $m_B \sim 5$ GeV
Lifetime of B hadrons:

$c\tau \sim 470\ \mu m$ (mixture $B^+/B^0/B_s$) , $\sim 390\ \mu m$ (Λ_b)
for E(B) ~ 50 GeV, flight length ~ 5 mm, $d_0 \sim 500\ \mu m$

➔ Spatial tagging:
- Signed impact parameter of tracks (or significance)
- Secondary vertex



➔ Soft-lepton tagging:
- Low p_T electron from B (D)
- Low p_T muon from B (D)

(limited by Br: around 20% each)

Status of Inner Detector

Covers |η|<2.5, in 2T B-field

Granularity:

- pixels: 80M (50x400μm²)
- strips: 6.3M (~80μm, st. 40mrad)
- TRT straws: 400k (4mm)

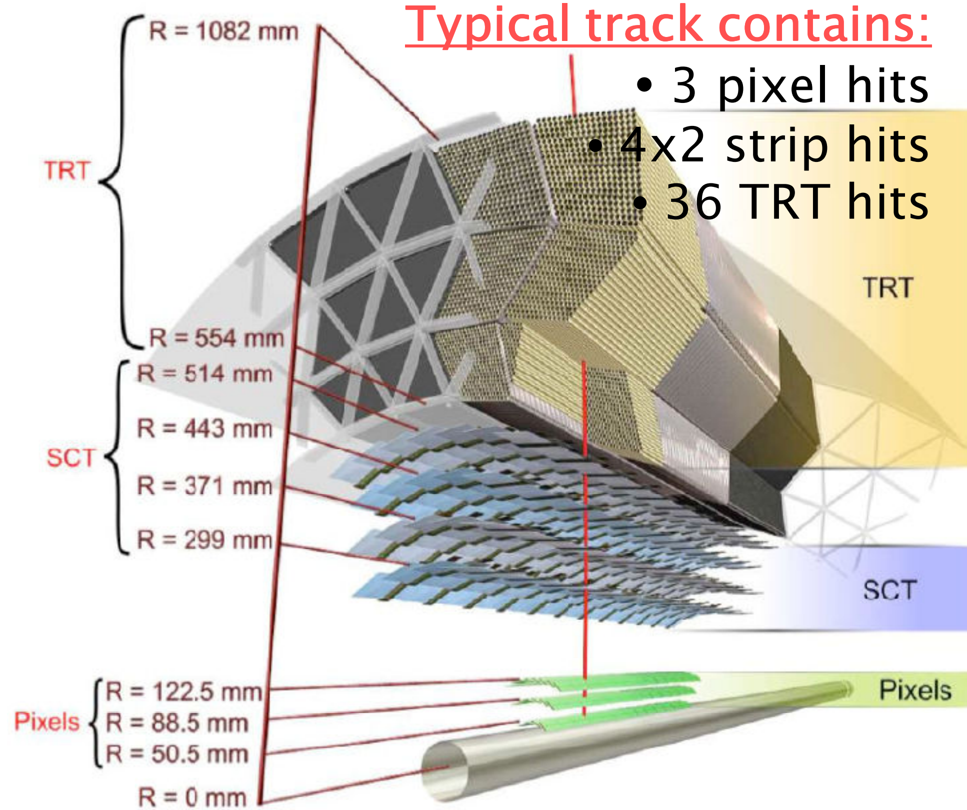
Active channels:

- Pixels: 98.0 %
- SCT: 99.3 %
- TRT: 98.2 %

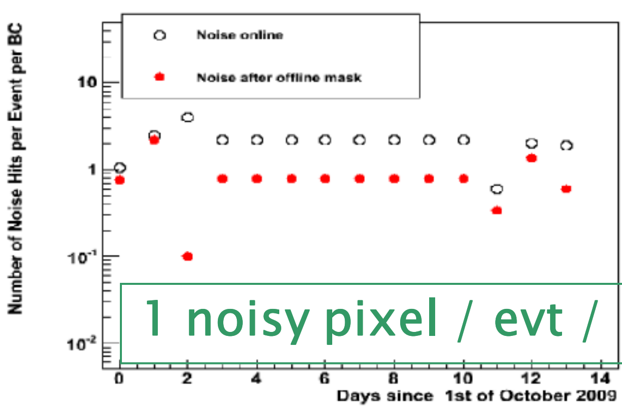
cf. Xin Wu's talk yesterday

Typical track contains:

- 3 pixel hits
- 4x2 strip hits
- 36 TRT hits



"Routine" monitoring plot:



1 noisy pixel / evt / BC !!

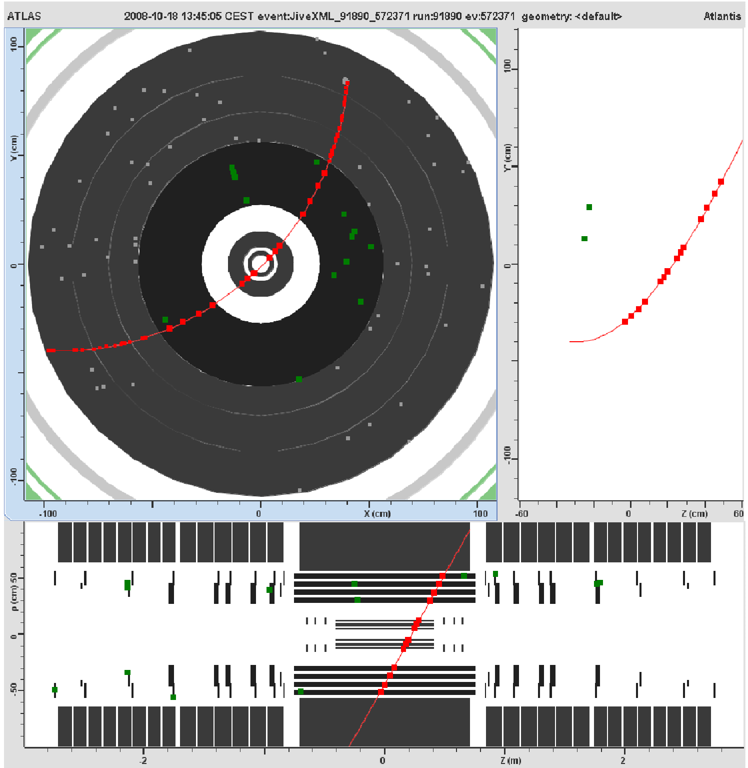
Noise occupancy: as expected

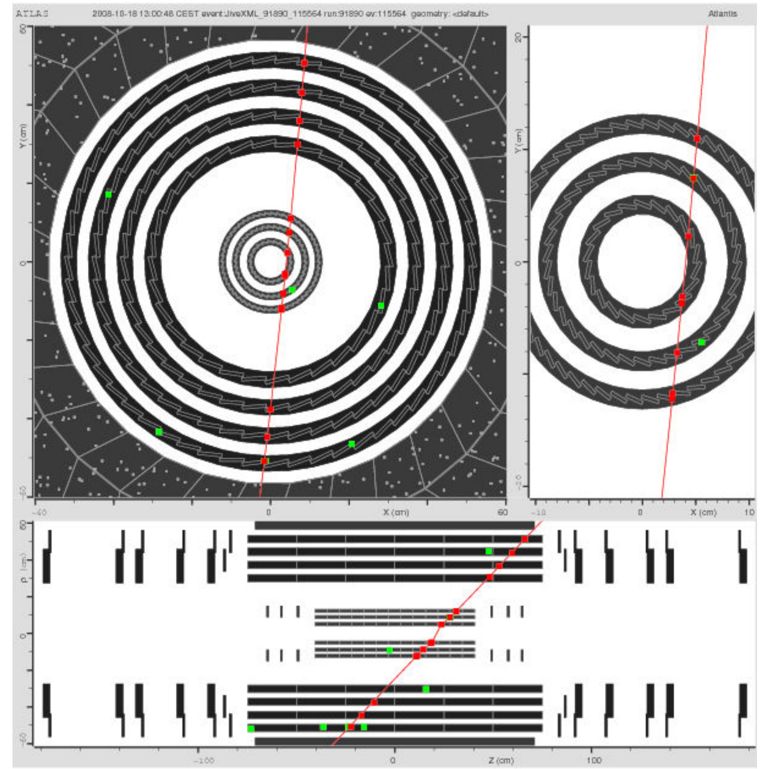
- Pixels: $\sim 10^{-10}$
- SCT: $\sim 10^{-5}$
- TRT: ~2 %

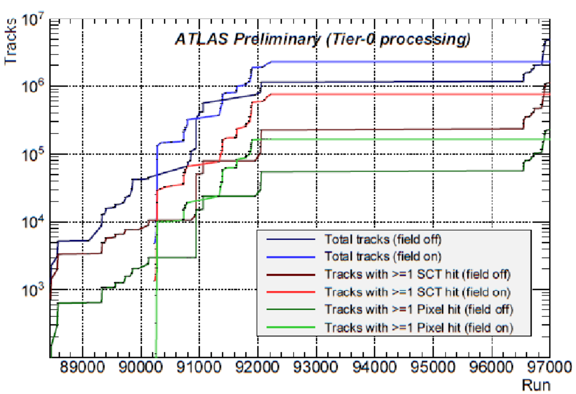
Timed-in, many measurements with cosmic rays since 2008 !

Cosmics data

cf. Oliver Kortner's talk yesterday





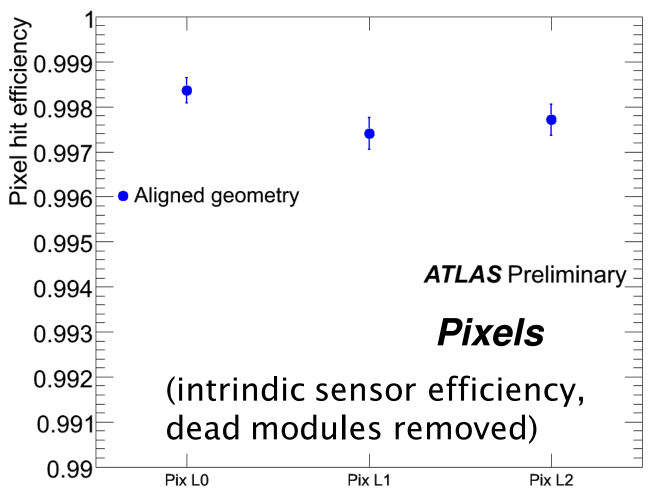


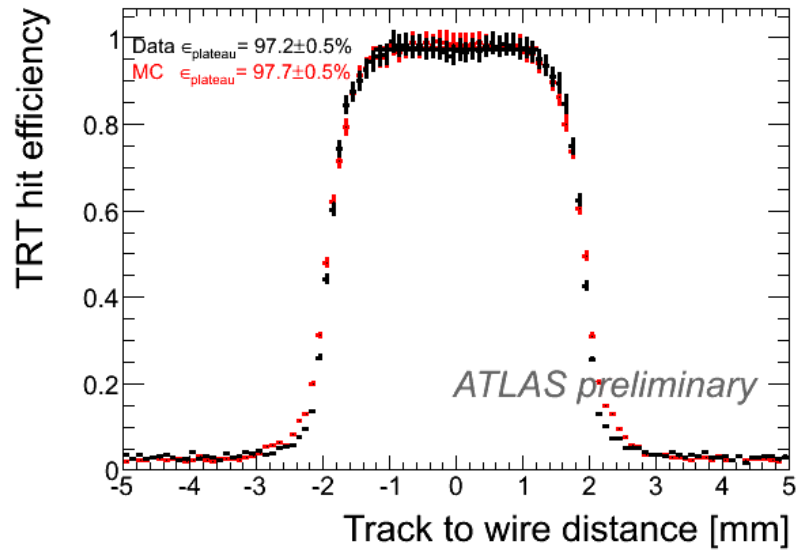
Autumn 2008:
- 7.6 M of tracks
- 230k (solenoid off) + 190k (on) crossing pixels

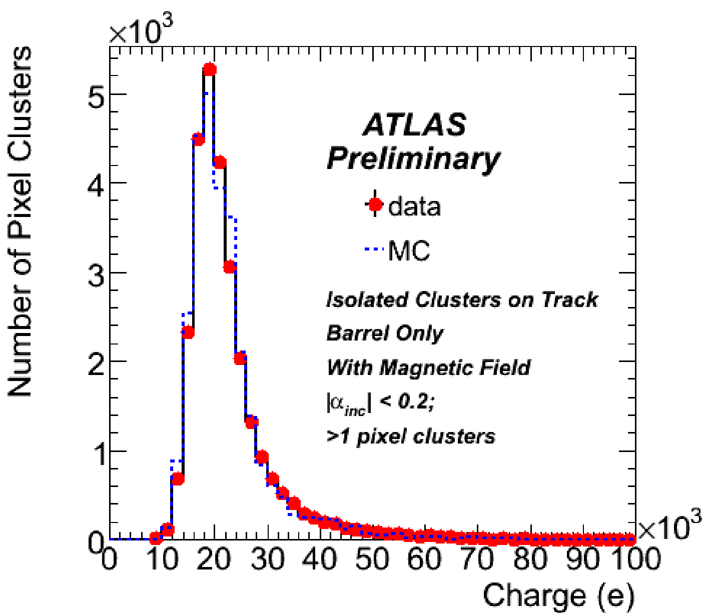
2009:
- currently accumulating statistics
- since Oct 1st : 390k (solenoid off) + 460k (on) tracks crossing full ID

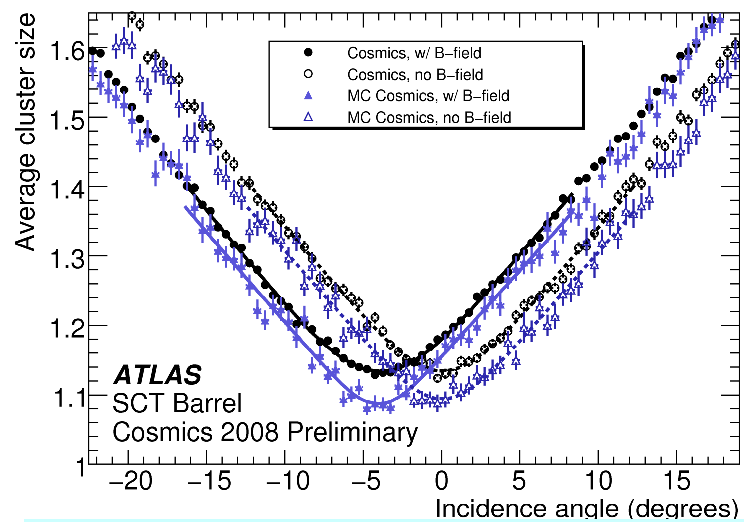
Basic ingredients for tracking

cf. Xin Wu's talk yesterday



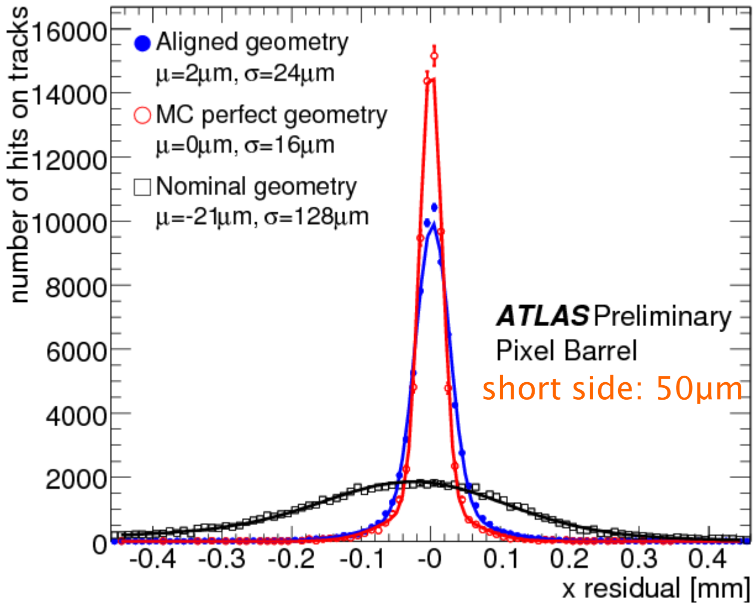


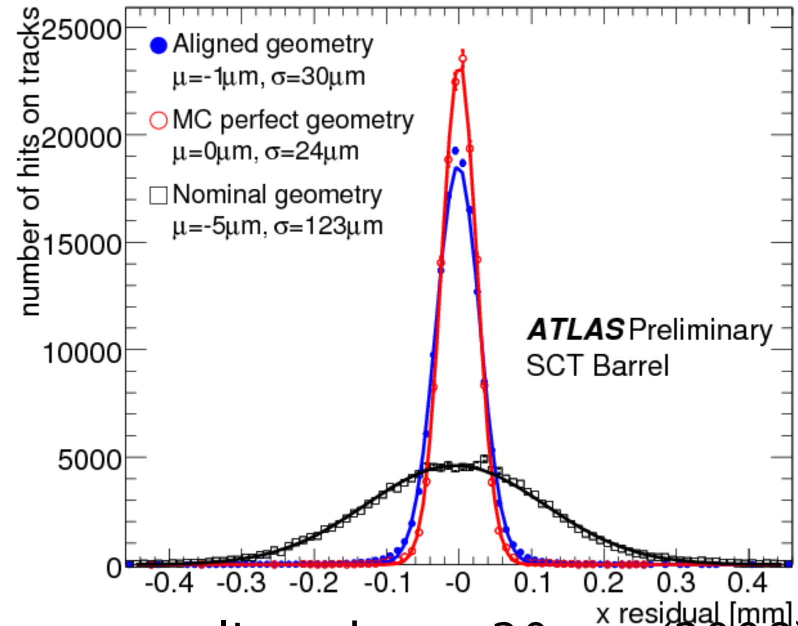


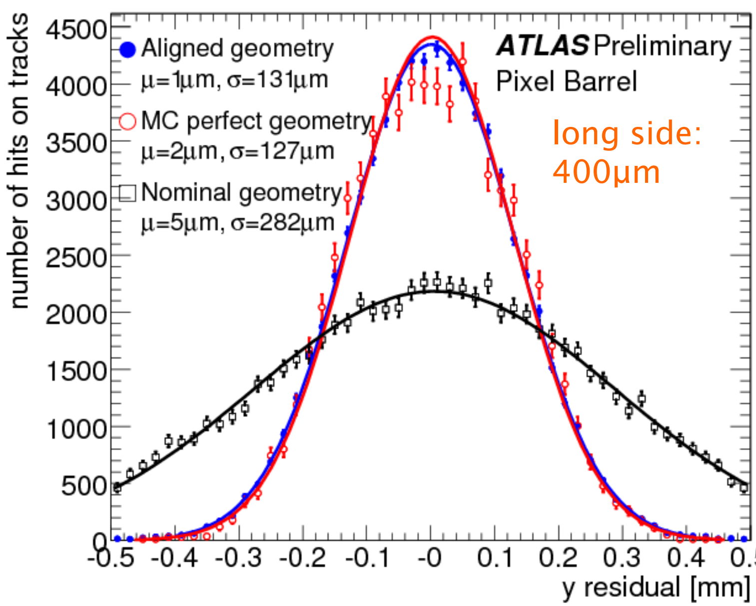


θ_L (data) = 3.93 ° ± 0.03 ° (stat) ± 0.10 ° (syst)
θ_L (model) = 3.69 ° ± 0.26 °

Alignment & cosmic muons: unbiased residuals

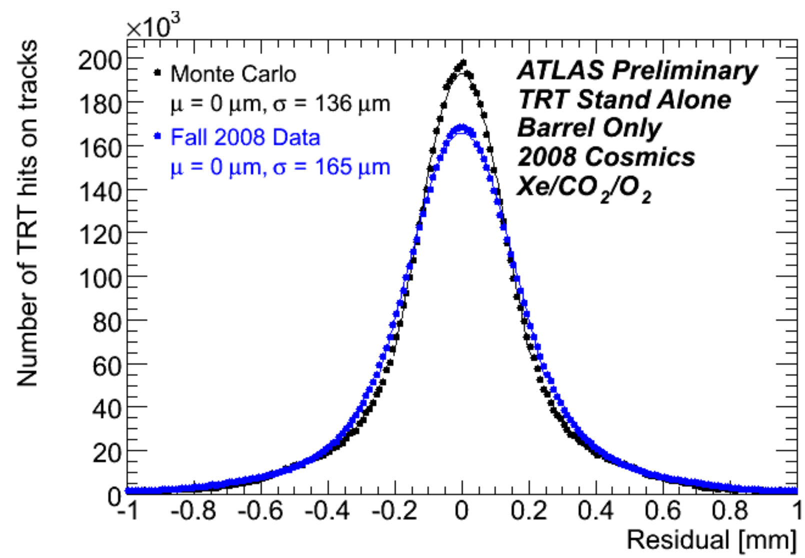






➔Si detectors aligned at ≲ 20 µm (2008)

➔stable in 2009

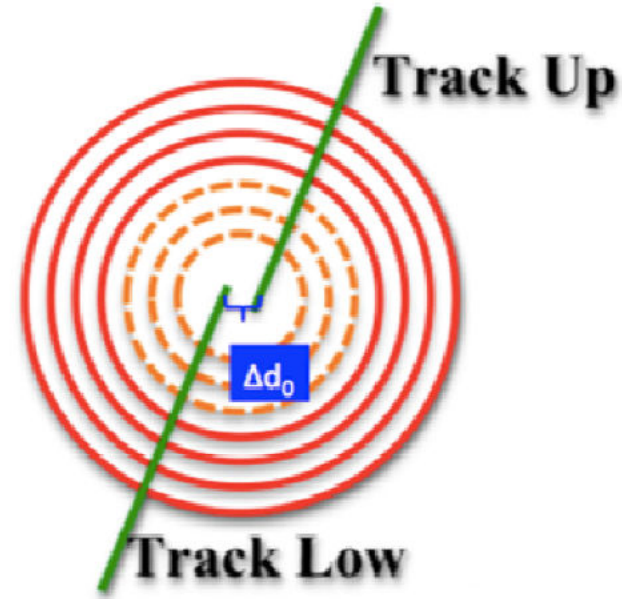


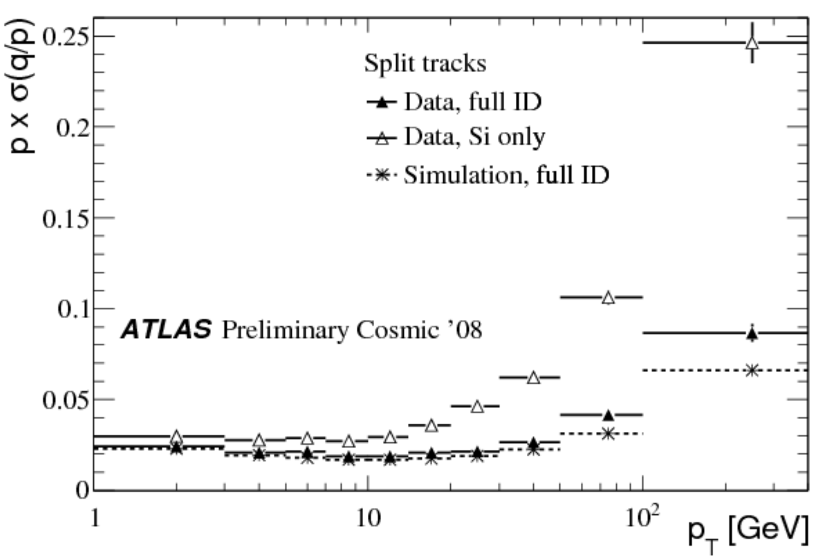
Resolution with Cosmics

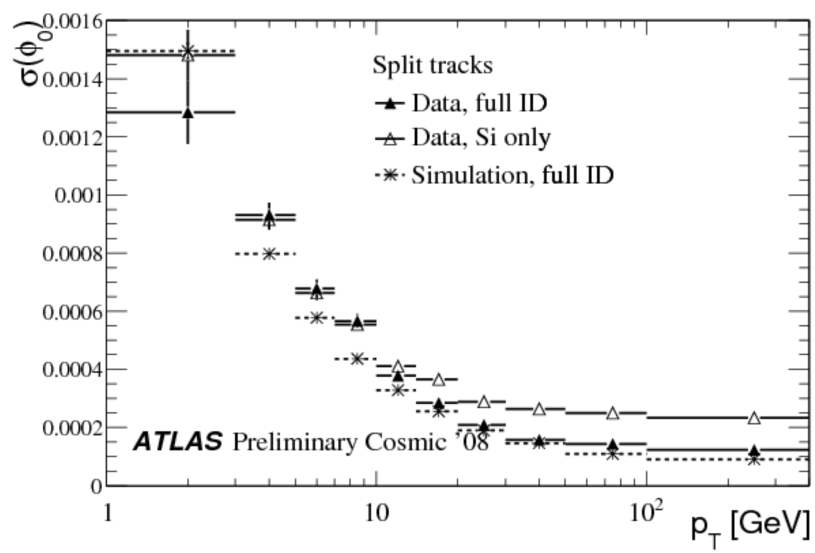
Split cosmic tracks into 2 collision-like tracks.
Refit the 2 tracks.
Quality cuts:

- Pt > 1 GeV/C
- Barrel hits only
- At least 2, 6 and 25 hits resp. in Pixels, SCT and TRT.
- |d0| < 40 mm

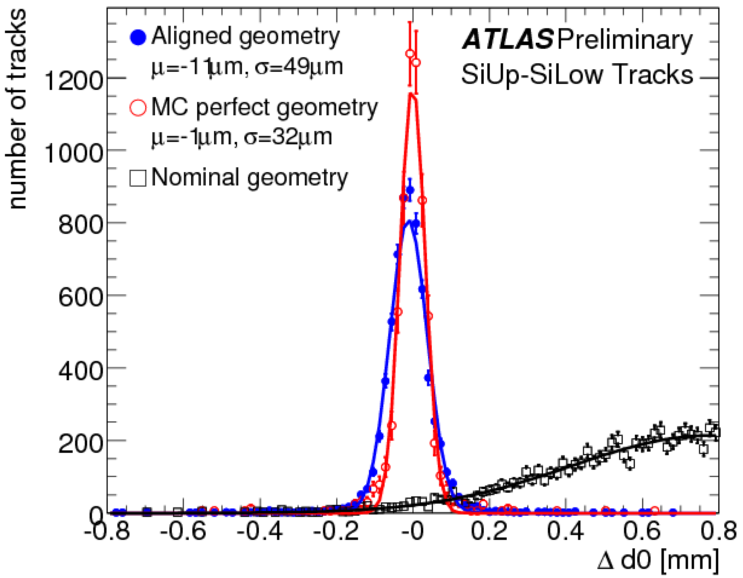
Look at the difference between the parameters of the 2 tracks.
Resolution = RMS of the residual / $\sqrt{2}$.

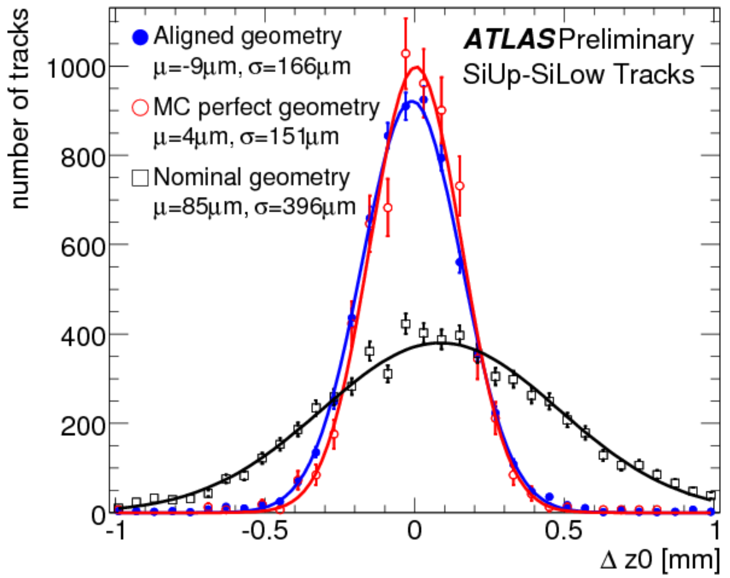


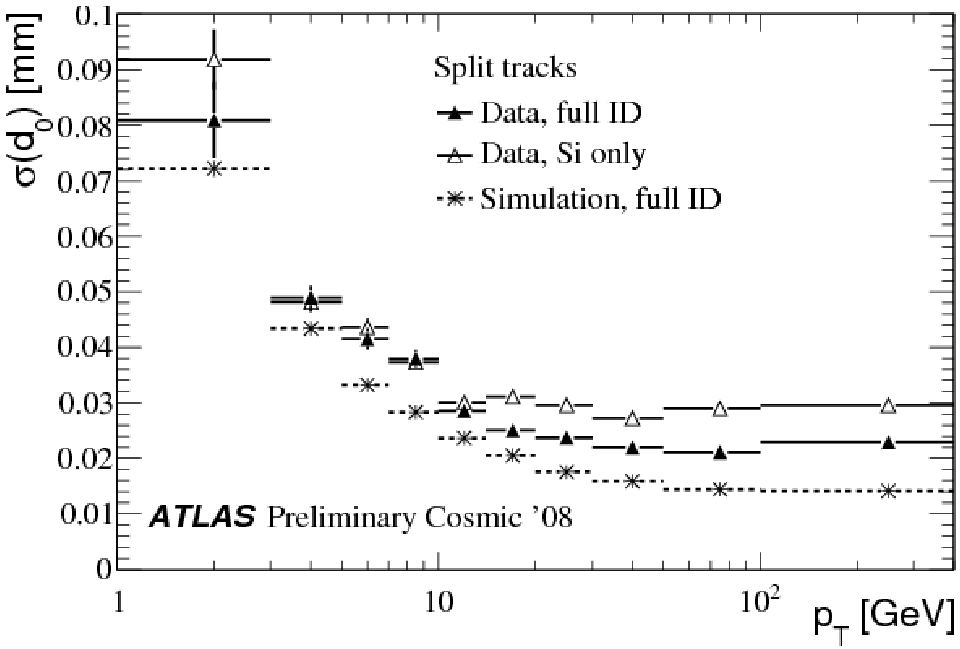




Impact parameter resolution





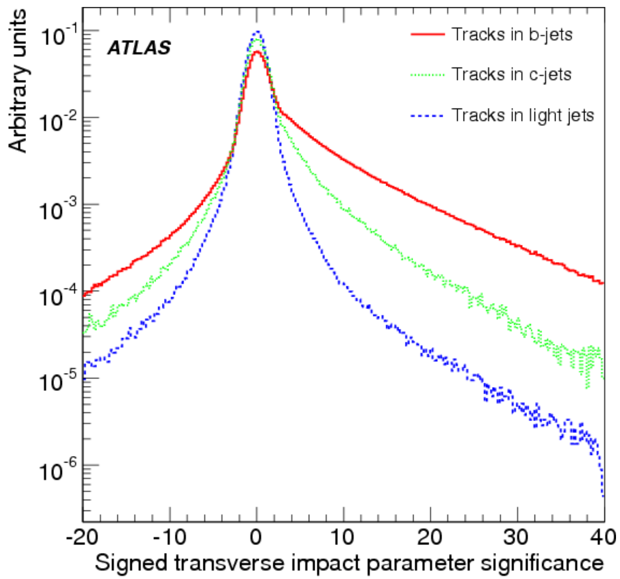


Typical track in a top b-jet:
$-p_T \sim 4$ GeV/c

Transverse impact parameter resolution:

- 44 μm expected from MC
- 48 μm measured

Early algorithms for b-tagging: overview

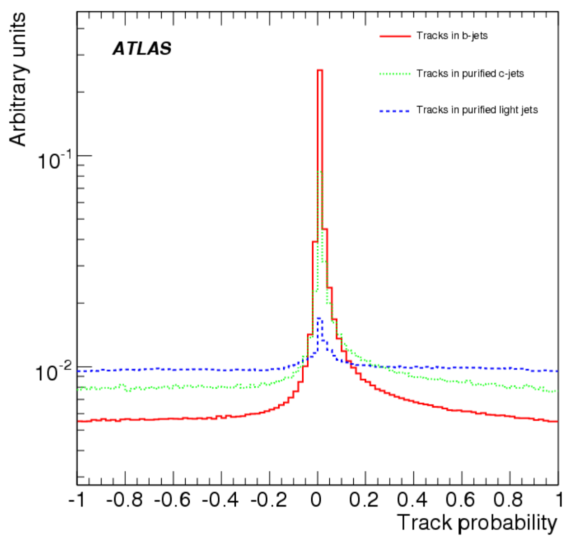


➔Relying on transverse impact parameter:

- ➔ TrackCounting: # of tracks with large d_0/σ
- ➔ JetProb: measuring compatibility of tracks with primary vertex, using a resolution function derived from data: it can be derived already with the 900 GeV data.

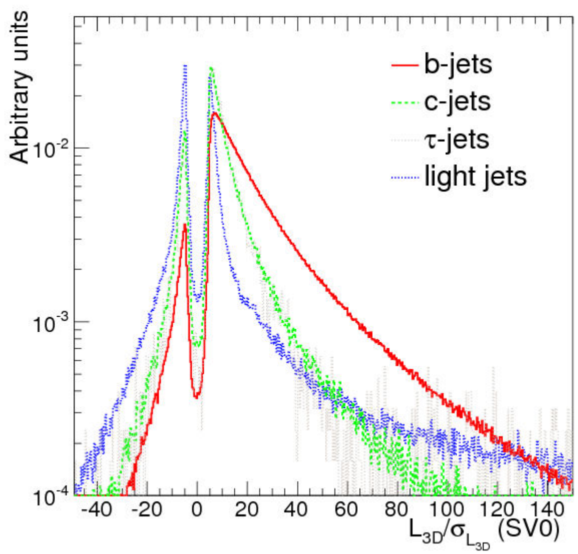
➔Relying on secondary vertex:

- inclusive secondary vertex



Track compatibility with primary vertex

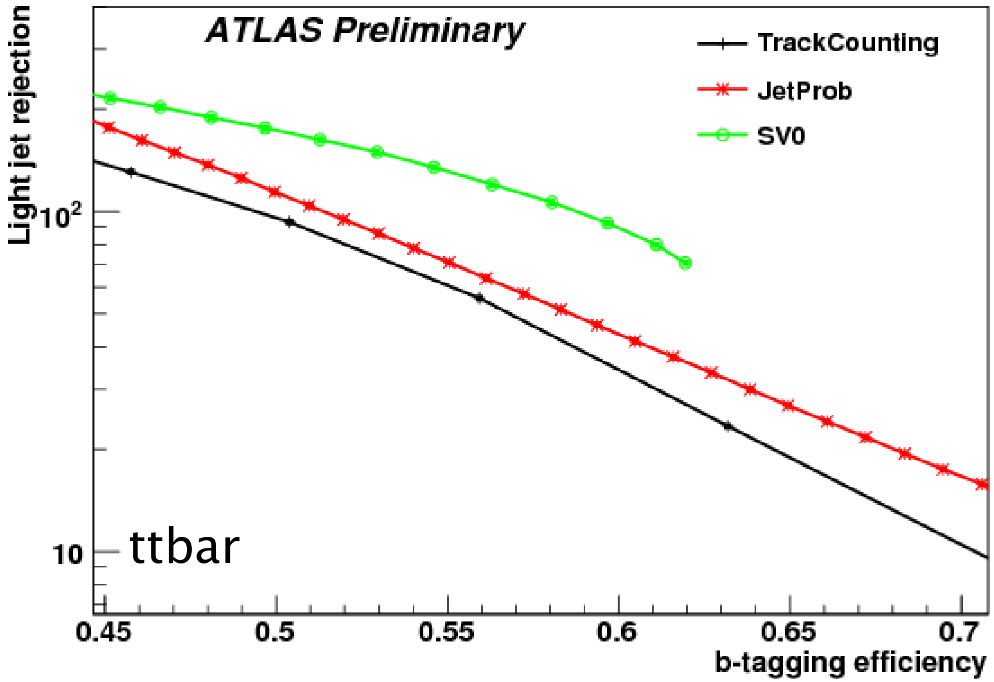
Normalized distance PV-SV



Early taggers: expected performance

Test sample:
- 500k ttbar events (10 TeV)
- rather central jets
- average p_T:
  - 70 GeV for b-jets
  - 55 GeV for light jets
- selection: p_T>15 GeV, $|\eta|<2.5$

Estimators: light jet rejection (inverse of mis-tagging rate) vs b-tagging efficiency.



	ϵ_b = 50%	ϵ_b = 60%
TrackCounting	96	38
JetProb	114	44
SV0	173	89

(errors stat.: ±1)

residual misalignment: not yet fold in, see later

Key ingredients

Track selection:

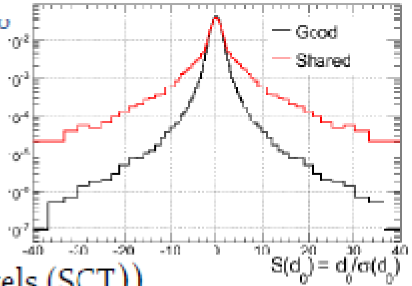
b-tagging relies on tracks reconstructed in the ID

- Selection optimized to reject :
  - "fake tracks"
  - tracks from long-lifetime particles (Kaons, Lambda, etc.)
  - tracks from material interactions (photons conversion and hadronic interactions)

- Selection cuts used for b-tagging :
  - $|p_T| > 1$ GeV
  - $|\eta| < 2.5$ } Detector acceptance
  - $|d_0| < 1$ mm
  - $|z_0| < 1.5$ mm
  - # hits b-layer : Nbla ≥ 1 → Good σ_{d0}
  - # hits in pixels : Npix ≥ 2
  - # hits in silicon (Pixel+SCT) ≥ 7 } Reject fakes

  Rejects long-lived particles (V°) and secondary interactions

+explicit V0 reco/rejection



- Tracks categories :
  - good
  - shared (tracks sharing a least 1 (2) hit(s) with another track in pixels (SCT))
  - other possible criteria : #hits, hits quality (holes, ganged pixels, etc.)

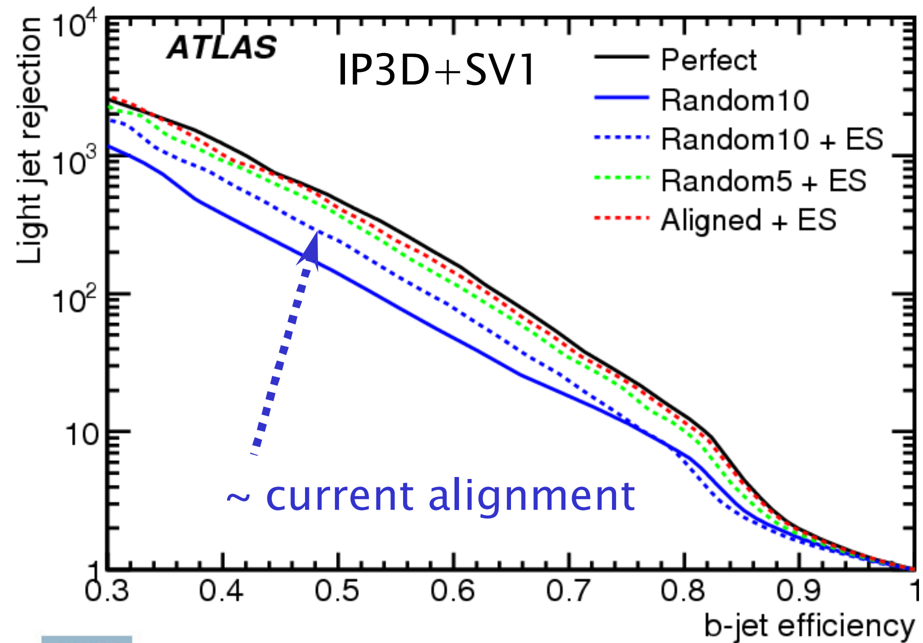
- With collisions @900 GeV : understanding tracking, MC validation with data

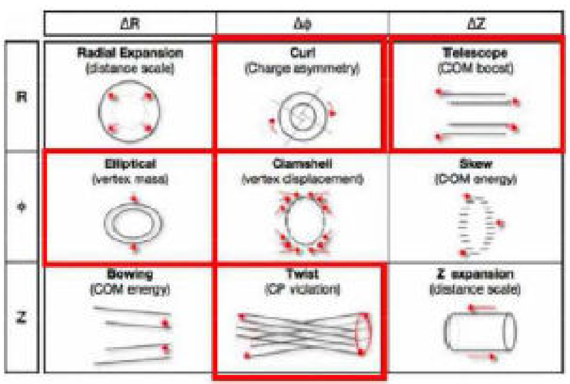
Impact of matter: MS, secondary interactions: 15% decrease in rejection when adding $0.02X_0$ (~10%) in silicon tracker.

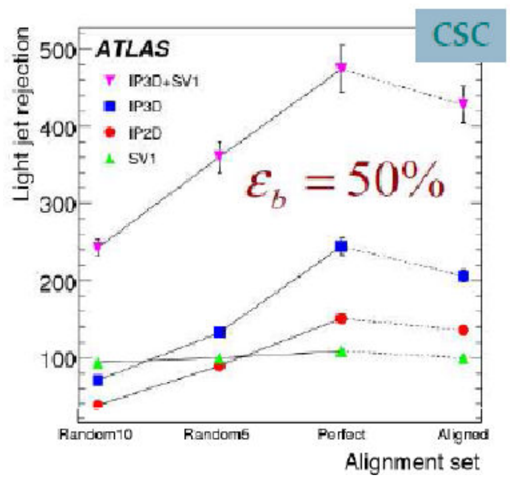
Impact of residual misalignments on b-tagging

MC geometry includes misplacements as expected in detector (incl. surveys)
10–100 μm for modules, mm for elements, some systematic deformations (global shifts, clocking effects,…)

Several MC studies:
- realign (cheat) + additional residual misalignments
- full ATLAS alignment procedure
- impact of weak modes





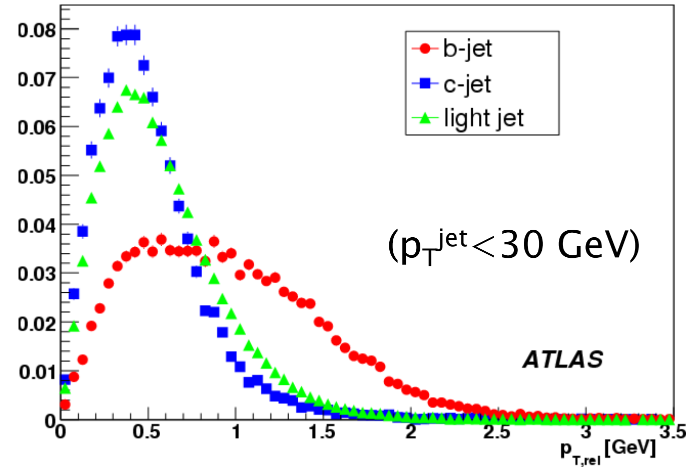


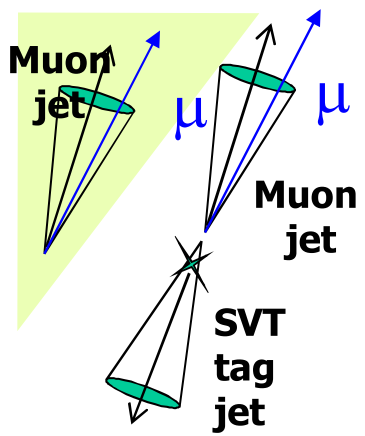
➔after full alignment, relative loss in rejection between 11-18%: encouraging, even though probably not all systematic deformations were accounted for

➔current actual alignment: x2 degradation (?), good enough for first data

Calibrating b-tagging efficiency with di-jet events

- Key-ingredient: soft muons
- Dedicated μ-jet trigger:
  - staged jet E_T thresholds
  - 1 Hz → 100k in 30h (1 pb^{-1}@10^{31} $cm^{-2}s^{-1}$)
- Method 1: p_T^{rel} templates
  - b & c templates from MC
  - fit in p_T, η bins



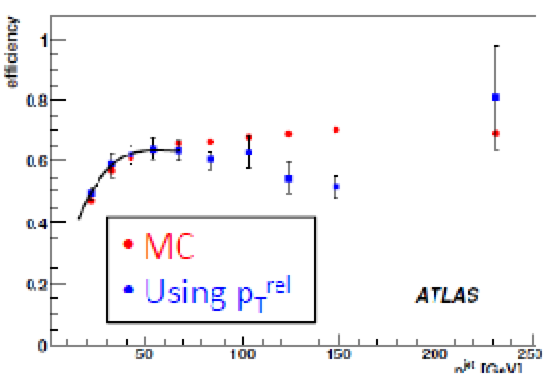


- Method 2: non-linear system (à la D0)
  - 2 samples
  - 2 different b-jet/light jet fractions
  - 2 non-correlated taggers
  - → system can be solved analytically

- Methods work well for p_T^{jet} < 80 GeV
- Syst. uncertainties dominate for > 50 pb^{-1}

→ Abs. precision on b-tag efficiency: 6%

→ Work in progress for mistag rates



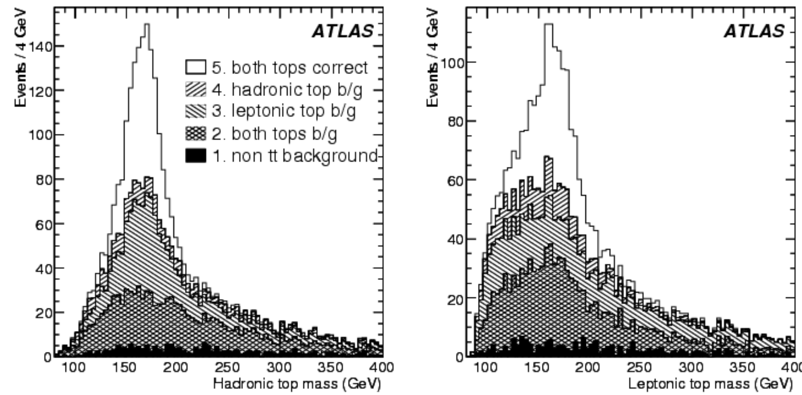
Calibrating b-tagging efficiency with top events

Tag counting method:

- count 1,2,3 b-tags → likelihood
- best accuracy on ε_b: for 100 pb^{-1}
  $\pm 2.7\%$(stat)$\pm 3.4\%$(syst)
- integrated efficiency only

Several selection methods:
- topological
- likelihood
- kinematic fit

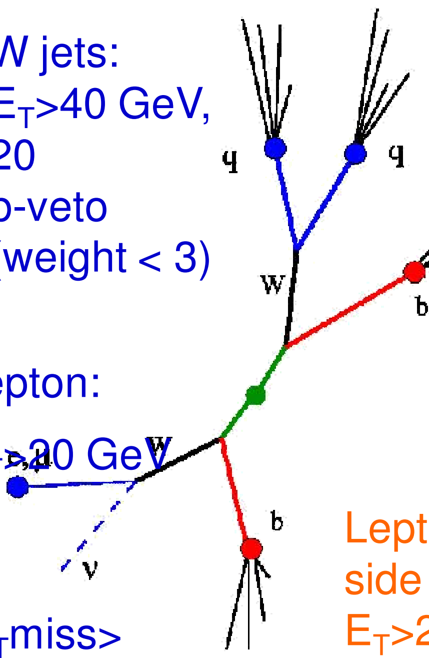
→ signal purity 60–80%



Selecting unbiased b-jets (leptonic side)

W jets:
E_T>40 GeV, 20
b-veto
(weight < 3)

Hadronic side
b-jet: E_T>40 GeV
b-tag
(weight>3)



Lepton:
E_T>20 GeV

$E_T$miss> 20 GeV

Leptonic side b-jet:
E_T>20 GeV
No tag requirement

b-tagging relative efficiency determination:

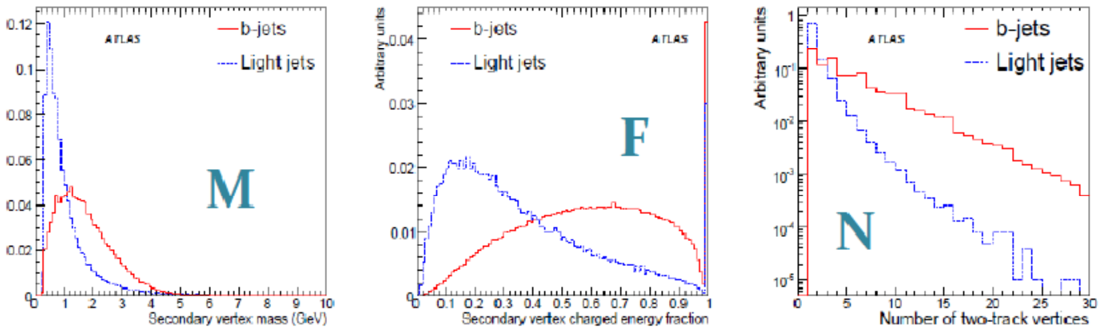
	Topo.	Like.	Kine.
Stat. (200 pb^{-1})	6.4%	4.4%	5.5%
Syst. error	3.4%	14.2%	6.2%

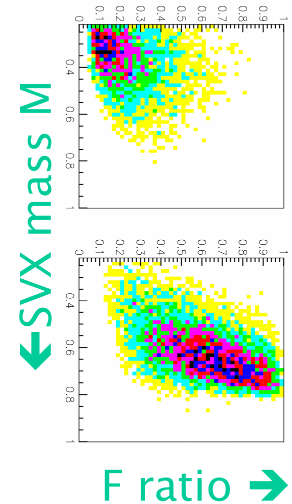
➔ can also be used to extract b’s p.d.fs

NB: studied @ 14 TeV

Beyond 2010: high-performance algorithms

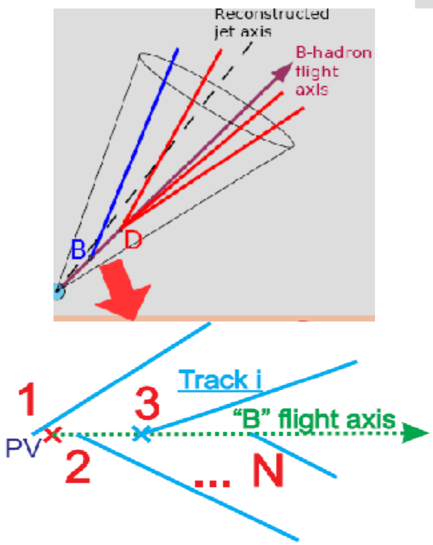
Likelihood ratio-based taggers: b- and light hypotheses



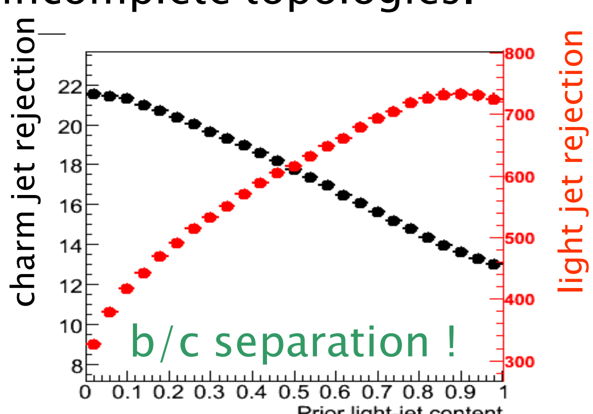


JetFitter: dedicated Kalman filter

$\{x_v, y_v, z_v, \phi, \theta, dist_1, dist_2, dist_3, \ldots, dist_N\}$



constrain all tracks from B/D hadron decays to intersect same flight axis → treatment of incomplete topologies.



Performance: light jet rejection

	$\epsilon_b = 50\%$	$\epsilon_b = 60\%$
SV0	173	89
IP3D	287	74
IP3D+SV1	1050	286

(ttbar events)

Conclusion

- ATLAS Inner Detector is working extremely well
- Important knowledge about its performance gained with cosmic data
- Allowed also to perform initial alignment (mostly in barrel)

  ➔ Most parameters close to nominal (MC expectations) values

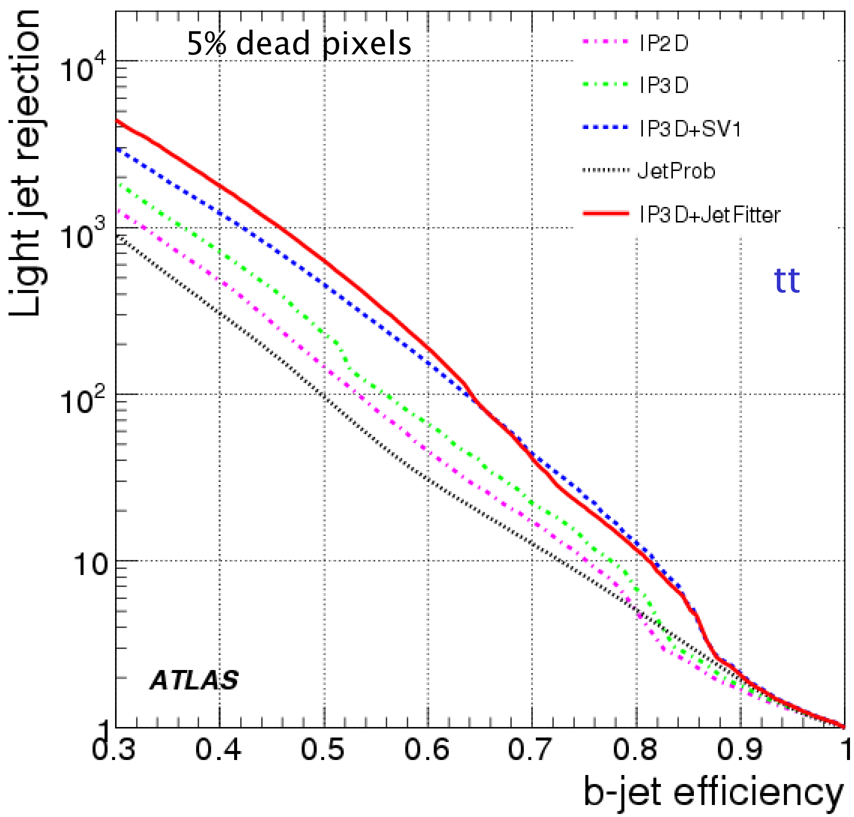
- Based on those results, some confidence that b-tagging could be operational very quickly with first data
- Specific simple and robust tagging algorithms designed for early data: should reach a light jet rejection of ~100 for ϵ_b=60% (~50 right at the beginning)
- It is furthermore very important to commission quickly (but thoroughly) b-tagging for top rediscovery at 7 TeV
  - Will start with tracking commissioning and MC comparison already at 900 GeV
  - Performance can be measured in data with jet events (50 pb^{-1}): efficiency (±6%) & mistags

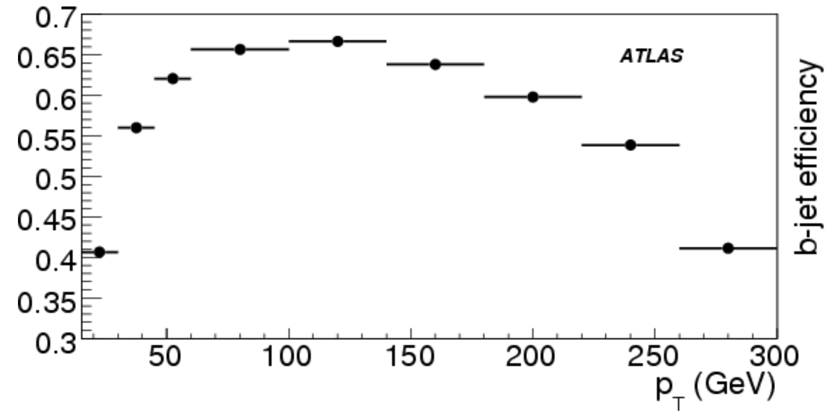
➔ Looking forward to first collisions to complete the commissioning work and get ready for LHC physics !

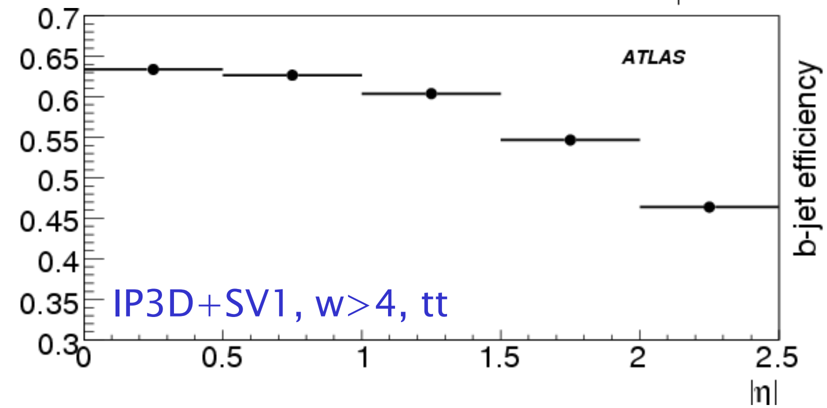
Back-up slides

Performance of b-tagging algorithms (CSC book)

Early taggers not available/optimized yet







- Track counting: R~30 @ 60%
- Soft muon: R~300 @ 10% (i.e. 80% w/ BR)
- Soft electron: R~100 @ 8%
- HLT: R~20 @ 60%
- Charm rejection: 5 to 7 @ 60%, up to 20 with JetFitter

Factorization (≠ channels): dependency on jet pT, η and env. (ΔR)

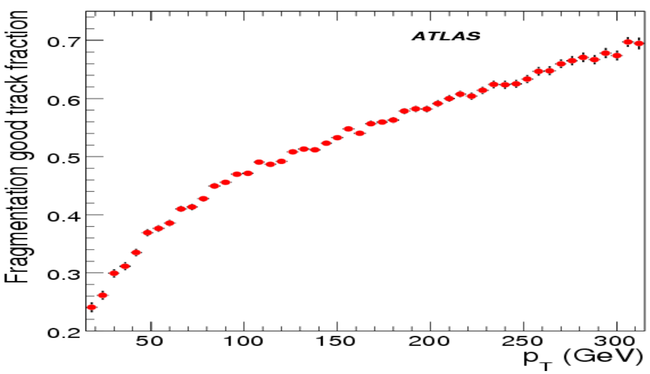
Degraded performance:

- low p_T: MS, secondaries
- high η: lever-arm (z_0),secondaries
- high p_T: next slide

Tagging high and very high-p_T jets

Challenges:

- fixed ΔR for track/jet assoc → dilution for jets of high p_T



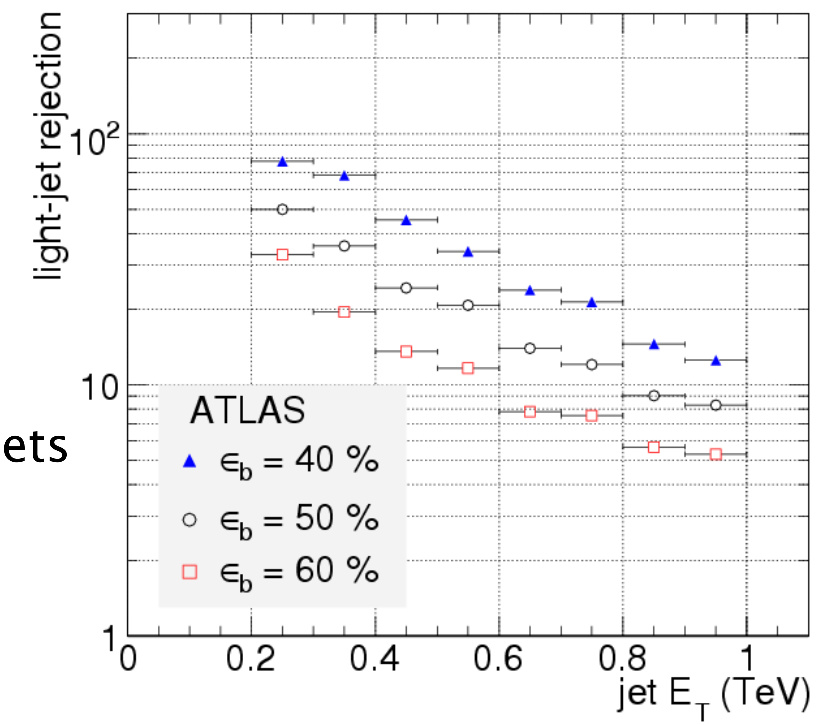
- gains (x2,x3) possible for non-isolated jets
- pattern-recognition issues
- 'late' B-decays

Fraction of jets (WH400):

	$R_B > 2.9$ cm	$R_B > 5.1$ cm
all E_T	9.0%	2.8%
$E_T > 100$ GeV	12.2%	3.9%
$E_T > 200$ GeV	21.1%	7.9%

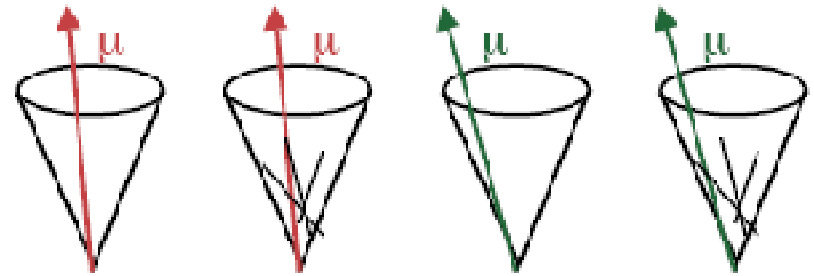
Example: Z'(2 TeV) → bb, uu

after retuning of IP3D+SV1:
(x3 worse otherwise)

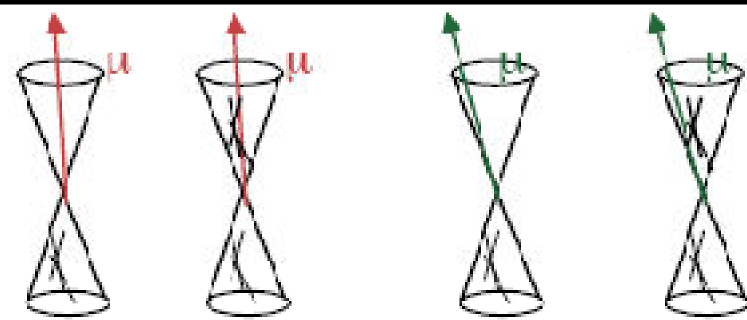


→ Require dedicated treatment for clustering, pattern-recognition

System8 on di-jet events



MJ	MJ LT	MJ MT	MJ LMT
8252	3985	2865	1681



DT	DT LT	DT MT	DT LMT
1501	803	555	349

Measurements

$$n = n_b + n_{cl}$$

$$p = p_b + p_c$$

$$n_\mu = \varepsilon^\mu n_b + r^\mu n_{cl}$$

$$p_\mu = \varepsilon^\mu p_b + r^\mu p_{cl}$$

$$n_{Tr} = \varepsilon^{Tr} n_b + r^{Tr} n_{cl}$$

$$p_{Tr} = \beta \varepsilon^{Tr} p_b + \alpha r^{Tr} p_{cl}$$

$$n_{all} = k_b \varepsilon^\mu \varepsilon^{Tr} n_b + k_{cl} r^\mu r^{Tr} n_{cl}$$

$$p_{all} = k_b \beta \varepsilon^\mu \varepsilon^{Tr} p_b + k_{cl} \alpha r^\mu r^{Tr} p_{cl}$$

→ system solvable analytically for ε, r and sample composition !

k: correlation between soft muon & track-based taggers

α,β: sample dependency

Tag counting in top events

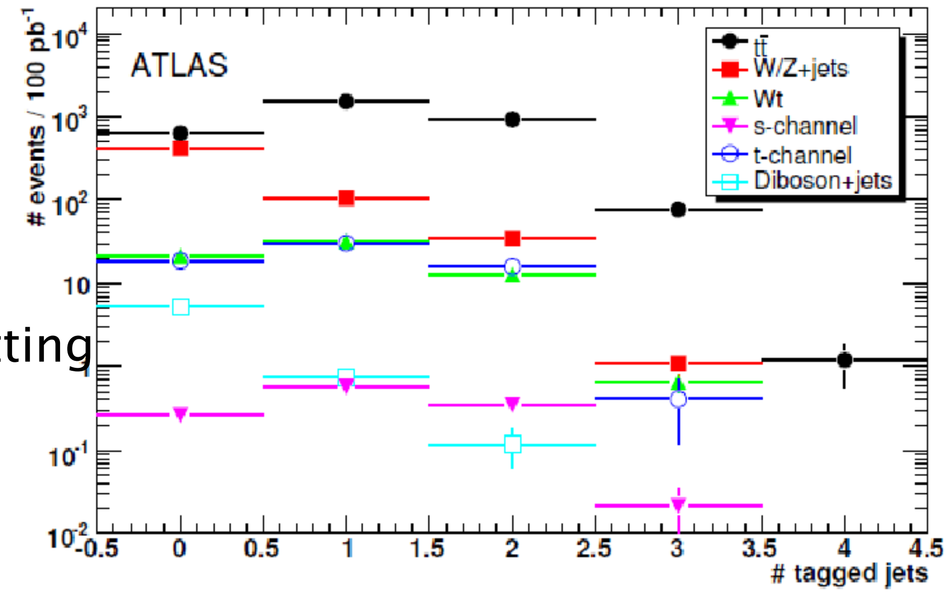
- selection ttbar l+jets: 4j > 20 GeV
- counting events with 1,2,3 tags
- HF contents:
  - 2 b-jets
  - possibly 1 c-jet from W
  - possibly cc/bb from gluon splitting
- first order:

$$N_{1tag} = N.2\varepsilon_b.(1-\varepsilon_b)$$

$$N_{2tags} = N.\varepsilon_b{}^2$$

- actual: (complex!) likelihood with rough estimate of R_u

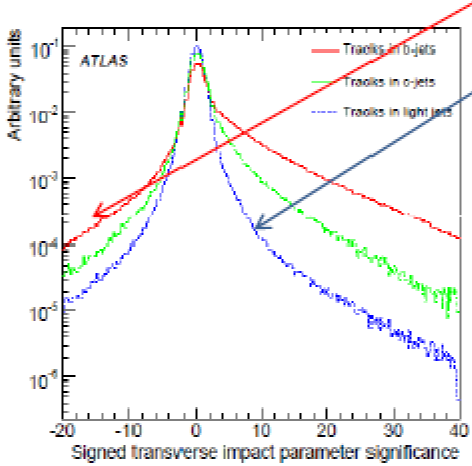
→4% total error with 100 pb^{-1}
→slightly better for di-lepton



Systematics	eps(b)	eps(c)	sigma(tt)
Light & Tau Jets (+-100%)	<0.1	38	<0.1
b-jet Labelling	1,4	12	0.1
Correlation between tags	<0.2	<0.2	<0.2
Jet Energy Scale (5%)	0.3	2,5	7
b-Jet Energy Scale (1%)	<0.1	1,5	0.8
(MC stat.	0.5	7	0.5)
Background (+-100%)	2,9	1,8	4,7
(AcerMC vs MC@NLO	0.5	13	8)
ISR/FSR	1,3	10	9
Top mass (+-2 GeV)	+0.5/-0.1	-	2,2
Total	**3,5**	**43**	**13**
Stat. (100 pb^{-1})	**2,2**	**16**	**1,8**

Calibrating mistag rates with jet events

- **Mistag rate** = fraction of light jets tagged (using a positive tagger ↔tagger using only tracks with $d_o>o$) $\varepsilon_l^{>0}$ = **mistag rate**
- Using jet events



Assumptions :

- Same contribution of tracks from b and light jets on $d_o<o$ side $\varepsilon_b^{<0} = \varepsilon_l^{<0}$
- Symmetrical distribution for tracks in light jets $\varepsilon_l^{>0} = \varepsilon_l^{<0}$

- Use negative tag method (tagger using only tracks with $d_o<o$)

N : **Number of jets in sample (b and light jets)**

f_b : **fraction of b jets**

$N^{<0}$: **Number of jets tagged using the negative tagger**

$$N^{<0} = f_b N \varepsilon_b^{<0} + (1 - f_b) N \varepsilon_l^{<0}$$
$$\Leftrightarrow N^{<0} = f_b N \varepsilon_l^{<0} + (1 - f_b) N \varepsilon_l^{<0}$$
$$\Leftrightarrow N^{<0} = N \varepsilon_l^{<0} = N \varepsilon_l^{>0}$$
$$\Leftrightarrow \varepsilon_l^{>0} = \frac{N^{<0}}{N}$$

Corrections using MC (or templates) :

- Tails for light jets on $d_o>o$ side (due to V°s, γ conversions and material interactions)
- Larger tails for b-jets on $d_o<o$ side (cascade decays)

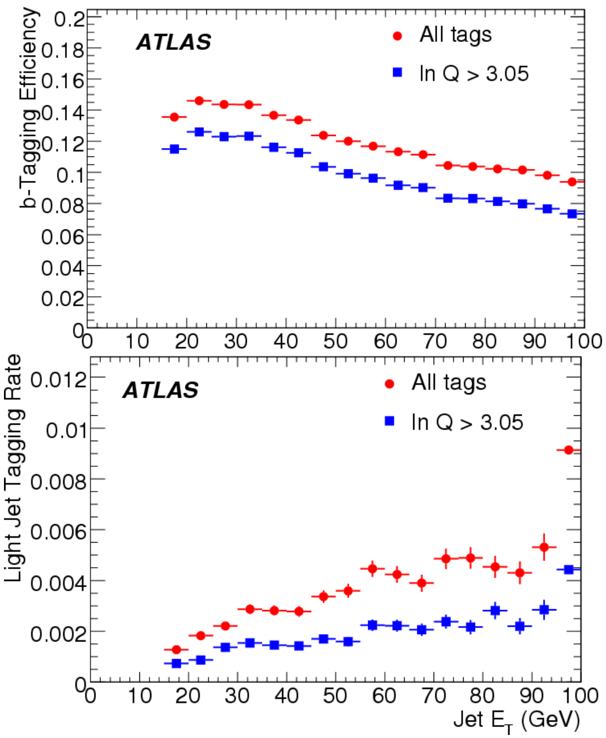
Very preliminary results not released yet…

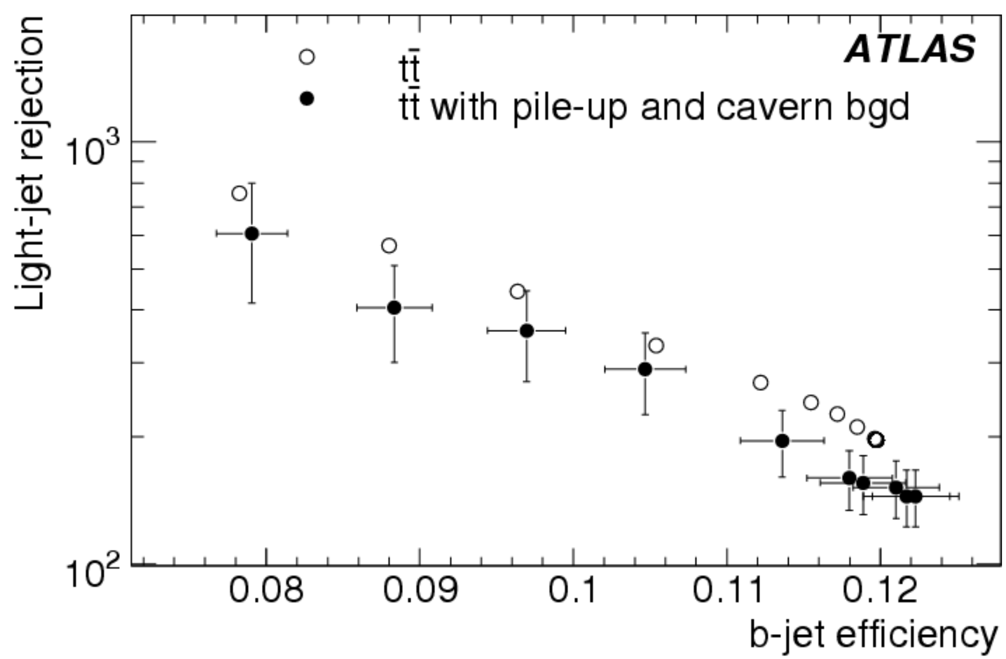
Tagging with Soft Muons

Br(b→μνX)+Br(b→c→μνX) = 11% + 10%

Muon reconstruction: high-efficiency down to 4 GeV/c p_T (90% above 5 GeV), low fakes (5p1000)

2 steps: association muon-jet (cone $\Delta R < 0.5$), likelihood ratios (1D, 2D) for b/light hypothesis using pT, pTrel.





➔ Rejection of 300 for 10% b-tagging efficiency (incl. Br)

➔ Fake muons vs pile-up/cavern bckgd: 15% impact on rej. (@ 2.10^{33})

Impact of misalignments: CSC challenge

- Misalignment set at three levels:
  - L1: barrel/end cap, L2: layer/disc L3: module

	LEVEL 1	X	Y	Z	α	β	γ
TRT	TRT Barrel	+1	+1	+1	0.20	-0.05	0
	TRT Endcap A	+2	-1	+2	-0.15	0.10	0
	TRT Endcap C	-2	+2	-3	-0.20	-0.15	0
SCT	SCT Barrel	+0.70	+1.20	+1.30	0.10	0.05	0.80
	SCT Endcap A	+2.10	-0.80	+1.80	-0.25	0.00	-0.50
	SCT Endcap C	-1.90	+2.00	-3.10	-0.10	0.05	0.40
Pixel	Whole	+0.60	+1.05	+1.15	-0.10	0.25	0.65

(displacements in mm; rotations in mrad)

L2	Layer	Systematic radial shift	Random shift in X,Y
TRT	Layer 0	+1.0 mm	0.2 mm
	Layer 1	-0.5 mm	0.1 mm
	Layer 2	+1.5 mm	0.3 mm

LEVEL2

	Layer/Disk	X	Y	Z	α	β	γ
Pixel Barrel	0	0.020	0.010	0	0	0	0.006
	1	-0.030	0.030	0	0	0	0.005
	2	-0.020	0.030	0	0	0	0.004
SCT Barrel	0	0	0	0	0	0	-0.001
	1	0.050	0.040	0	0	0	0.009
	2	0.070	0.080	0	0	0	0.008
	3	0.100	0.090	0	0	0	0.007
SCT Endcap A	1	0.050	0.040	0	0	0	-0.001
	2	0.010	-0.080	0	0	0	0
	3	-0.050	0.020	0	0	0	0.001
	4	-0.080	0.060	0	0	0	0.002
	5	0.040	0.040	0	0	0	0.003
	6	-0.050	0.030	0	0	0	0.004
	7	-0.030	-0.020	0	0	0	0.005
	8	0.060	0.030	0	0	0	0.006
	9	0.080	-0.050	0	0	0	0.007
SCT Endcap C	1	0.050	-0.050	0	0	0	0.008
	2	0	0.080	0	0	0	0
	3	0.020	0.010	0	0	0	0.001
	4	0.040	-0.080	0	0	0	-0.008
	5	0	0.030	0	0	0	0.003
	6	0.010	0.030	0	0	0	-0.004
	7	0	-0.060	0	0	0	0.004
	8	0.030	0.030	0	0	0	0.006
	9	0.040	0.050	0	0	0	-0.007

LEVEL3

	x	y	z	α	β	γ
Pixel Barrel modules	0.030	0.030	0.050	0.001	0.001	0.001
Pixel Endcap modules	0.030	0.030	0.050	0.001	0.001	0.001
SCT Barrel modules	0.150	0.150	0.150	0.001	0.001	0.001
SCT Endcap modules	0.100	0.150	0.150	0.001	0.001	0.001

- Shifts are realistic !
  - Though may seem huge
  - Surveyed during assembly…

Alignment stability